June 1974

ISIS: AN INTERACTIVE FACILITY FOR SCENE ANALYSIS RESEARCH

by

J. M. Tenenbaum
T. D. Garvey
S. A. Weyl
H. C. Wolf

Artificial Intelligence Center

Technical Note 95

SRI Project 1187

<u>Proceedings Second International Joint Conference on Pattern Recognition</u>, Lyngby-Copenhagen, Denmark, August 1974.

The research reported herein was primarily sponsored by the Office of Naval Research under Contract N00014-71-C-0294. Additional support was provided by the Advanced Research Projects Agency under Contract DAHC04-72-C-0008, and NASA Contract NASW-2086.

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **JUN 1974**	2. REPORT TYPE	3. DATES COVERED **00-00-1974 to 00-00-1974**
4. TITLE AND SUBTITLE **ISIS: An Interactive Facility for Scene Analysis Research**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **SRI International,Artificial Intelligence Center,333 Ravenswood Avenue ,Menlo Park,CA,94025-3493**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR’S ACRONYM(S)
		11. SPONSOR/MONITOR’S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**		
13. SUPPLEMENTARY NOTES		

14. ABSTRACT

This paper summarizes initial progress in developing a computer system that can be rapidly programmed to analyze any class of pictorial scenes. Scene annlysis programs have been awkward to develop using conventional programming systems because of the difficulty of formulating pictorial descriptions in symbolic terms. Picture processing techniques are inherently ad hoc and must be deduced empirically for each application. We have constructed an interactive system specifically designed for expressing and experimenting with perceptual strategies. ’!:’he system allows an experimenter to describe basic perceptual concepts to a computer in terms of pictorial examples. The examples are designated graphically by encircling areas of a displayed scene with a cursor. A concept is represented internally by values of primitive feature-extraction operators that distingUish it from examples of previously defined concepts. Concepts so defined constitute a common vocabulary, shared by man and machine, that can be used symbolically in describing objects and specifying scene analysis procedures. The system bas been used to formulate interactively descriptions that distinguish objects 1n indoor room scenes and programs that locate these objects 1n images~

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**4**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

ISIS: AN INTERACTIVE FACILITY FOR SCENE ANALYSIS RESEARCH

J. M. Tenenbaum, T. D. Garvey, S. A. Weyl, H. C. Wolf
Stanford Research Institute
Menlo Park, California

Abstract

This paper summarizes initial progress in developing a computer system that can be rapidly programmed to analyze any class of pictorial scenes. Scene analysis programs have been awkward to develop using conventional programming systems because of the difficulty of formulating pictorial descriptions in symbolic terms. Picture processing techniques are inherently ad hoc and must be deduced empirically for each application.

We have constructed an interactive system specifically designed for expressing and experimenting with perceptual strategies. The system allows an experimenter to describe basic perceptual concepts to a computer in terms of pictorial examples. The examples are designated graphically by encircling areas of a displayed scene with a cursor. A concept is represented internally by values of primitive feature-extraction operators that distinguish it from examples of previously defined concepts. Concepts so defined constitute a common vocabulary, shared by man and machine, that can be used symbolically in describing objects and specifying scene analysis procedures.

The system has been used to formulate interactively descriptions that distinguish objects in indoor room scenes and programs that locate these objects in images.

A Description of ISIS

ISIS, Interactive Scene Interpretation System, is an integrated set of INTERLISP functions facilitating the development and testing of pictorial representations. The system consists primarily of an image file, a library of primitive feature extraction functions, a means for applying selected primitive operators to graphically designated areas of a scene, a symbolic data structure "for accumulating concept definitions, and an iconic data structure" for retaining pictorial examples of those concepts. A detailed description of the system with applications is available from the authors (SRI Artificial Intelligence Center, Technical Note 87).

The symbolic data structure is a semantic net whose nodes represent prototype patterns for objects, attributes, and relations. Prototypes are defined symbolically in terms of other nodes or iconically by reference to designated regions of a digitized image. The system can obtain values for unspecified properties of a concept by applying operators to its examples. Figure 1(a-c) illustrates the semantic net. In Figure 1(c), the object TABLETOP and the attribute BUFF are defined iconically, while the object FLOOR and the attribute HORIZONTAL are defined symbolically in terms of other nodes. FLOOR is distinguished from TABLETOP by height.

Iconic regions are represented as an explicit list of picture samples and/or as a list of vertices describing a closed polygonal boundary. Efficient routines exist for determining whether a given picture element is contained within a bounded region, for obtaining a set of samples over a bounded region, and for fitting a boundary around a set of samples.

The primitive functions and symbolic data structure reside within the interactive environment of INTERLISP (formerly BBN-LISP) on a PDP-10. The raw data arrays, and support routines for graphics, file handling, coordinate transformations, etc., reside in a FORTRAN environment accessible as a lower fork through the TENEX operating system.

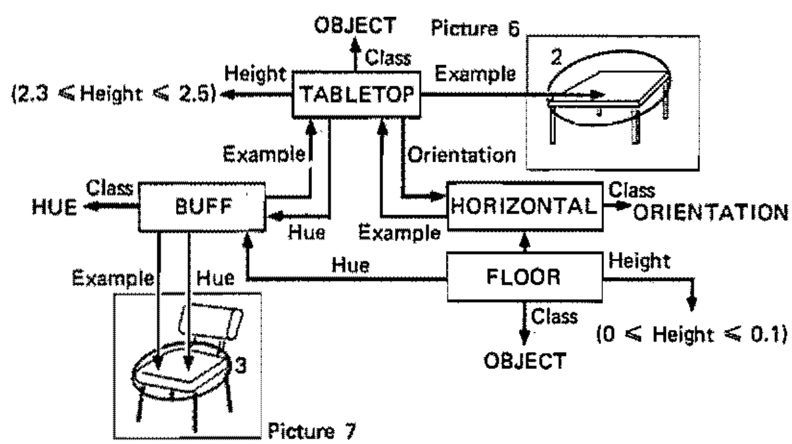
Images are stored in 120 X 120 sample arrays, each sample characterized by TV brightness through red, green, blue, and neutral filters and by a range measurement. (These range measurements are the simulated output of a time-of-flight range finder, currently under development.) A high resolution vector display and an MOS-refreshed color video monitor with a low resolution vector overlay serve as graphic output devices. Both displays allow users to select points on the screen using a cursor.

NAME	
CLASS	
PROPERTIES	(Values of operators legal for this class of nodes.)
IS IT?	(A procedure to determine whether something is an example of this concept.)
FIND A!	(A procedure to find an example of this concept.)
EXAMPLES	(A list of known examples of this concept.)

(a) STRUCTURE OF A GENERAL NODE

NAME — Tabletop	
CLASS — Object	
PROPERTIES —	HUE: Buff ORIENTATION: Horizontal
IS IT?	Procedure to test whether an object has the distinguishing features of Tabletop.
FIND A!	Procedure to find a Tabletop in a given picture.
EXAMPLES	Picture 6, Region 2

(b) EXAMPLE OF A NODE



(c) LINKAGE OF CONCEPTS AND EXAMPLES

SA-1187-1R

FIGURE 1 THE SEMANTIC NET

The Use of ISIS

To describe concepts in terms of the computer's primitive functions, a human must know (or be able to determine easily) what those primitive functions can distinguish. Our system provides the trainer with tools with which to chose primitive functions empirically. The system first displays a digitized scene as a color image or as a thresholded gradient. This display qualitatively conveys which color boundaries are easily discriminated by the computer. Users can circle regions of the displayed image and obtain from the system either average or extreme numerical values for local operators, such as height, hue, saturation, and surface normal. Thus, by applying operators to examples and counterexamples of pictorial concepts in the image, the trainer can discover directly which operators provide sufficient discrimination. Users can try out a proposed description by requesting the system to illuminate all parts of the displayed scene that correspond to the description.

The use of ISIS to formulate pictorial concepts is best illustrated by example. The following examples are based on the simple room scene shown in Figure 2.

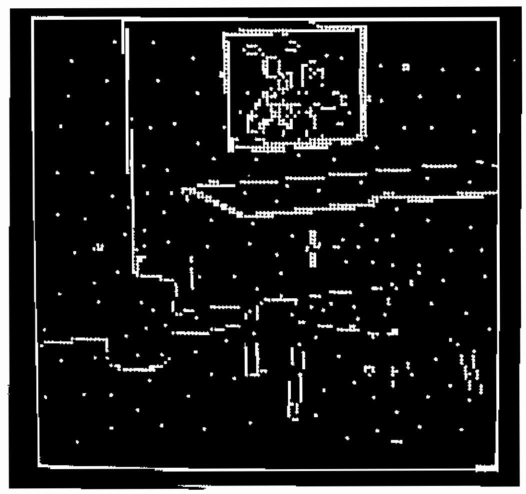
SA-1187-23

FIGURE 2 DIGITIZED ROOM SCENE VIEWED ON COLOR MONITOR

Description of Attributes

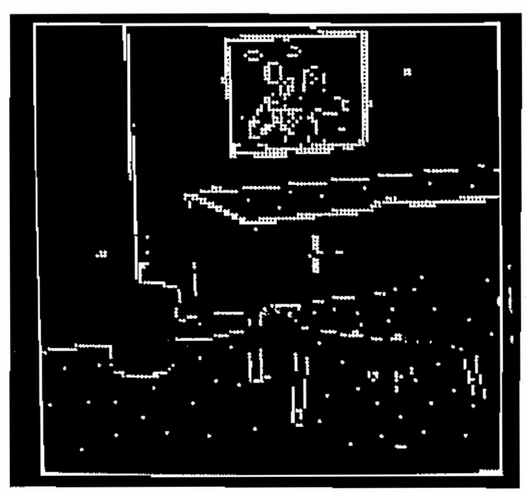
Using ISIS, the trainer can define common surface attributes, such as the color BUFF and the orientation HORIZONTAL, with which to describe surfaces symbolically. These attribute labels can usually be defined by a characteristic range of values for a single primitive operator. The range of values can be determined empirically by applying the operator to pictorial examples of the attribute. HORIZONTAL, for instance, might be defined as any surface whose normal is within 5 degrees of the Z axis based on values of the ISIS function ORIENT applied to image samples on the FLOOR and TABLETOP.

The adequacy of a proposed description can be tested by requesting the system to intensify points in the displayed image that satisfy a proposed description. This test is accomplished by using the description to filter random image samples. In Figure 3, a set of random image samples is shown superimposed on a gradient display of the office scene. In Figure 4, only points satisfying the suggested definition of HORIZONTAL are retained. Figure 4 validates the definition of HORIZONTAL, since all intensified points are on surfaces such as the FLOOR, TABLETOP, and CHAIRSEAT--normally thought of as HORIZONTAL.



SA-1187-7

FIGURE 3 RANDOM IMAGE SAMPLES



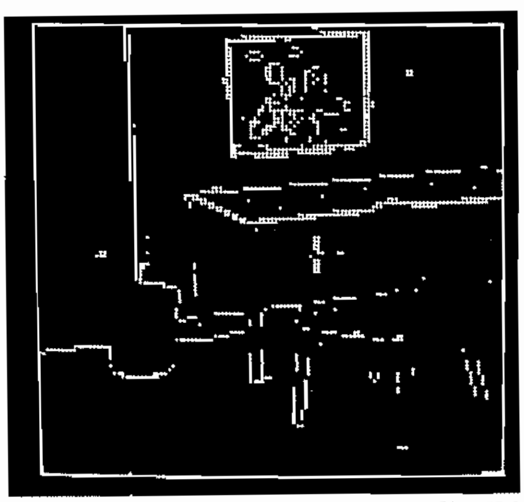
SA-1187-10

FIGURE 4 HORIZONTAL IMAGE SAMPLES

Description of Objects

Defined attributes can be used symbolically to describe objects. For example, TABLETOP might now be defined to the system as a HORIZONTAL surface distinguished by height. The appropriate height range was determined by applying the ISIS function HEIGHT to a few manually designated TABLETOP samples. Figure 5 confirms that height and orientation adequately distinguish TABLETOP in this particular scene.

It is not always possible to specify a simple predicate that will select all image samples belonging to a desired object and to no others. For such cases, the repertoire of ISIS primitives described in this summary can be augmented with special purpose feature extraction routines. Procedures which validate edges, grow regions and bound rectangular surfaces have been developed by several users and are now generally available on ISIS. Further examples of the use of ISIS for characterizing objects can be found in Technical Note 87 and in a paper by Garvey and Tenenbaum (see On the Automatic Generation of Programs for Locating Objects in Office Scenes, this volume).



SA-1187-9

FIGURE 5 RANDOM SAMPLES WITH HEIGHT AND SURFACE ORIENTATION OF TABLE TOP

Goals for Interactive Scene Analysis

Ideally, one would like to program a computer to find an unfamiliar object as one would instruct a person, by providing a crude description of the desired object. A tree, for example, might be described as a "green, leafy region above a tall, brown, vertical, bark-textured region." The computer could demonstrate comprehension by outlining instances of the described object in a displayed image. The programmer could then refine the description empirically to correct errors in the computer's interpretation.

Communication on this level with a machine is hampered because the machine does not share an "understanding" with the human of basic descriptive concepts like "green," "leafy," "above," "vertical," and "bark-textured." Such concepts are often difficult to express in natural language, and even more so in conventional programming languages. Our aim has therefore been to create an interactive system that allows an experimenter to describe basic perceptual concepts to the computer using pictorial examples. The examples are designated graphically by encircling portions of a displayed scene with a cursor. Given a pictorial example, a representation can be empirically constructed by trying pictorial operators in order of increasing complexity until the example is sufficiently distinguished from previously determined representations. Pictorially defined concepts constitute a shared vocabulary that can be used to describe objects or to define more complex concepts. This paradigm is intended to elevate the user above the details of hand-coded algorithms, allowing him instead to concentrate on the construction of descriptions and strategies.

Acknowledgments

This work evolved from formulative discussions with Martin Fischler, a strong and early advocate of an interactive approach to scene analysis. We also wish to acknowledge John Gaschnig who helped implement an earlier version of this system.

The research reported herein was primarily sponsored by the Office of Naval Research under Contract N00014-71-C-0294. Additional support was provided by the Advanced Research Projects Agency under Contract DAHC04-72-C-0008, and NASA Contract NASW-2086.